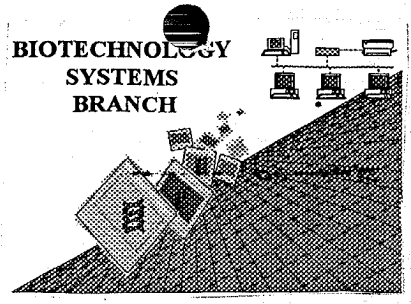
Gillespie

#8

BIOTECHNOLOGY SYSTEMS BRANCH



RECEIVED
NOV 13 2000
TECH CENTER 1600/2900

RAW SEQUENCE LISTING ERROR REPORT

RECEIVED
DEC 06 2000
TECH CENTER 1600/2900

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/508,821

Source: 1655

Date Processed by STIC: 10/27/2000

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin30help@uspto.gov or phone 703-306-4119 (R. Wax)

TO **REDUCE** ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

http://www.uspto.gov/web/offices/pac/checker

Raw Sequence Listing Error Summary

ERROR DETECTED **SUGGESTED CORRECTION** SERIAL NUMBER: 09/508,821

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

RECEIVED
NOV 13 2000
TECH CENTER 1600/2900

1 _____ Wrapped Nucleics
The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

2 _____ Wrapped Aminos
The amino acid number/text at the end of each line "wrapped " down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

3 _____ Incorrect Line Length
The rules require that a line not exceed 72 characters in length. This includes spaces.

4 _____ Misaligned Amino Acid Numbering
The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

RECEIVED
DEC 06 2000
TECH CENTER 1600/2900

5 _____ Non-ASCII
This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 _____ Variable Length
Sequence(s) ____ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.

7 _____ PatentIn ver. 2.0 "bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**

8 _____ Skipped Sequences (OLD RULES)
Sequence(s) ____ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 _____ Skipped Sequences (NEW RULES)
Sequence(s) ____ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

10 _____ Use of n's or Xaa's (NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11 _____ Use of <213>Organism (NEW RULES)
Sequence(s) _____ are missing this mandatory field or its response.

12 _____ Use of <220>Feature (NEW RULES)
Sequence(s) ____ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)

13 _____ PatentIn ver. 2.0 "bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other means to copy file to floppy disk.

J. Goldberg

1655

RAW SEQUENCE LISTING DATE: 10/27/2000
PATENT APPLICATION: US/09/508,821 TIME: 13:38:12

Input Set : A:\01294811.app
Output Set: N:\CRF3\10272000\I508821.raw

Does Not Comply
Corrected Diskette Needed

```
  3 <110> APPLICANT: ROULEAU, Guy A.
  4       JOOBER, Ridha
  5       BENKELFAT, Chawki
  7 <120> TITLE OF INVENTION: POLYMORPHIC CAG REPEAT-CONTAINING GENE AND USES THEREOF
  9 <130> FILE REFERENCE: 12948-10
 11 <140> CURRENT APPLICATION NUMBER: 09/508,821
 12 <141> CURRENT FILING DATE: 2000-05-26
 14 <150> PRIOR APPLICATION NUMBER: PCT/CA98/00884
 15 <151> PRIOR FILING DATE: 1998-09-18
 17 <150> PRIOR APPLICATION NUMBER: 2,216,057 CA
 18 <151> PRIOR FILING DATE: 1997-09-19
 20 <160> NUMBER OF SEQ ID NOS: 6
 22 <170> SOFTWARE: PatentIn Ver. 2.1
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 328
 26 <212> TYPE: DNA
 27 <213> ORGANISM: Mus musculus
 29 <400> SEQUENCE: 1
 30 agggcagcca ctttccccag cattcccagt ccttccctac ctcctccact tatgccccaa 60
 31 cagtgcaggg tggtgggcag ggggcccact cctacaagag ctgcacagca ccatctgccc 120
 32 agcctcatga taggccgatg agtgccaatg cgaacctggc tccagggcaa cgggtccaga 180
 33 atcttcacgc ttaccagcct ggccgccttg gctacgagca gcagcagcaa gcacttcaag 240
 34 gccgtcacca cacccaggaa acactccact accagaacct cgccaagtac caacactatg 300
 35 gacagcaagg ccagggctac tgtccacc                                    328
 38 <210> SEQ ID NO: 2
 39 <211> LENGTH: 330
 40 <212> TYPE: DNA
 41 <213> ORGANISM: Homo sapiens
 43 <400> SEQUENCE: 2
W--> 44 tccttcccca cctcctccac ctactcctcc tctgtcccag ggtggtgggc agggggccaa 60
 45 ctcctataag agttgcacag caccgactgc ccagccccat gacaggccgc tgactgccag 120
 46 ctccagcctg gccccggggc agcgggtcca gaatcttcat gcctaccagt cgggccgcct 180
 47 cagctatgac cagcagcagc agcagcagca gcagcagcag cagcagcagc aagcccttca 240
 48 gagccggcac catgcccagg aaaccctcca ttaccaaaac ctcgccaagt atcagcacta 300
 49 cgggcagcaa ggccagggct actgccagcc                                   330
 52 <210> SEQ ID NO: 3
 53 <211> LENGTH: 23
 54 <212> TYPE: DNA
 55 <213> ORGANISM: Artificial Sequence
 57 <220> FEATURE:
 58 <223> OTHER INFORMATION: Description of Artificial Sequence:
 59       oligonucleotide
 61 <400> SEQUENCE: 3
 62 ggggcagcgg gtccagaatc ttc                                        23
 65 <210> SEQ ID NO: 4
 66 <211> LENGTH: 23
 67 <212> TYPE: DNA
```

see item 10 on Error Summary Sheet

RAW SEQUENCE LISTING DATE: 10/27/2000
PATENT APPLICATION: US/09/508,821 TIME: 13:38:12

Input Set : A:\01294811.app
Output Set: N:\CRF3\10272000\I508821.raw

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence:
oligonucleotide
<400> SEQUENCE: 4
agcactacgg gcagcaaggc cag 23
<210> SEQ ID NO: 5
<211> LENGTH: 6022
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 5
ggatccagca ggcccaaggg gatgagggag cggaaattgc tctgctaaat gcttttgagc 60
tgtcaggaag ggctgggagt gatgggtggg ggacattggg gaggagctgg caatgggcgg 120
gggggggcg ggtagctccc cagtgacctg gcgctgggca gccggttttg cctcccgcat 180
cagtggccgt ccttggcaag actcagctgc aggcgatgtg ggagcgggaa ttacagagca 240
cacctccctg acacagaagt tgtcaatatg cgcacagctg gtggggaggc tcaggcgaag 300
gggggactat taagagctgc gcgggggagc aggcagggtg gggaggtggg tgggagggtg 360
ctttctgagg caaaaggaag tggcccgtct gaatcgctca tcctctgccc cctccctgcc 420
catcctcccc tccctccttc cctccctccc tcccttcctt tttcttttca cagataacca 480
gcccgagtca tgcagtcttt tcgagaaagg tgtggtttcc atggcaaaca acagaactac 540
cagcagacct cgcaggaaac atcacgccta gagaattaca ggcagccgag tcaggccggg 600
ctaagctgcg accggcagcg gctgctcgcc aaggactatt ataacccgca gccttacccg 660
agctatgagg gtggcgctgg cacgccctct ggcactgcag ccgcggtggc cgccgacaag 720
taccaccgag gcagcaaggc cctgcccaca cagcaaggcc tgcaggggag gccggctttc 780
cctggctacg gcgtccagga cagcagcccc tacccaggcc gctatgctgg tgaggagagc 840
cttcaggctt gggggcccc acagccacca ccccacagc cgcagccact acctgcaggg 900
gtggccaagt atgatgagaa cttgatgaaa aagacagcag tgcccccag caggcagtat 960
gcagagcagg gcgcccaggt gccctttcgg actcactccc tgcacgtcca gcagccaccg 1020
ccgccccagc agcccctggc ataccccaag ctccaaaggc agaagctgca gaacgacatt 1080
gcctcccctc tgcccttccc ccagggtacc cactttcctc agcattccca gtccttcccc 1140
acctcctcca cctactcctc ctctgtccag ggtggtgggc agggggccca ctcctataag 1200
agttgcacag caccgactgc ccagccccat gacaggccgc tgactgccag ctccagcctg 1260
gccccggggc agcgggtcca gaatcttcat gcctaccagt cgggccgcct cagctatgac 1320
cagcagcagc agcagcagca gcagcagcag cagcagcagc aagcccttca gagccggcac 1380
catgcccagg aaaccctcca ttaccaaaac ctcgccaagt atcagcacta cgggcagcaa 1440
ggccagggct actgccagcc ggacgcagcc gtccggaccc cagagcagta ctaccagacc 1500
ttcagcccca gctccagcca ctcacccgcc cgctccgtgg gccgctcacc ttcctacagt 1560
tccacaccgt cgccgctgat gccaaacctg gagaactttc cctacagcca gcagccgctc 1620
agcaccgggg ccttccccgc agggatcact gaccacagcc acttcatgcc cctgctcaat 1680
ccctccccaa cggatgccac cagctctgtg gacacccagg ctggcaactg caagcccctt 1740
cagaaggaca agctccctga gaacctgctg tcggatctca gcctgcagag cctcacggcg 1800
ctgaccttac aggtggagaa catctccaac accgtccagc agctgctgct ctccaaggct 1860
gctgtgccgc agaagaaagg tgtcaagaac ctcgtgtcca ggaccccaga gcagcataaa 1920
agccagcact gcagccccga agggagcggc tactcagccg agcccgcagg cacaccgctg 1980
tcagagccgc cgagcagcac gccacagtcc acgcatgcgg agccgcagga ggccgactac 2040
ctgagcggct ccgaggaccc actggagcgc agcttcctct actgcaacca ggcccgtggc 2100
agccctgcca gggtcaacag caactcgaag gccaagcccg agtccgtgtc cacctgttct 2160
gtgacctctc ctgacgacat gtccaccaaa tctgacgact ccttccagag cctacacggc 2220
agtctgccgc tcgacagctt ctccaagttc gtggcgggtg agcgggactg tccgcggctg 2280

RAW SEQUENCE LISTING DATE: 10/27/2000
PATENT APPLICATION: US/09/508,821 TIME: 13:38:12

Input Set : A:\01294811.app
Output Set: N:\CRF3\10272000\I508821.raw

122 ctgctcagcg ccctggcaca ggaggacctg gcctccgaga tcctggggct gcaggaagcc 2340
123 atcggtgaga aggccgacaa agcttgggct gaagcaccca gcctggtcaa ggacagcagc 2400
124 aagccaccct tctcgctgga gaaccacagc gcctgcctgg actctgtggc caagagtgcg 2460
125 tggccccggc ctggggagcc ggaggccctg cccgactcct tgcagctgga caagggcggc 2520
126 aatgccaagg acttcagccc agggctgttt gaagaccctt ccgtggcctt cgctacgcct 2580
127 gaccccaaaa agacaactgg tcctctctcc tttggtacca agcccaccct tggggttcct 2640
128 gctccagacc ccactacagc agcttttgac tgtttcccgg acacaaccgc tgccagctca 2700
129 gcggacagcg ccaaccccтt tgcctggcca gaggaaaacc tgggggatgc ttgtcccagg 2760
130 tggggattgc accctggcga gcttaccaag ggcctggagc agggtgggaa ggcctcagat 2820
131 ggcatcagca aaggggacac ccatgaggct tcggcctgcc tgggcttcca ggaggaggac 2880
132 ccccctgggg agaaggtggc ctcgttgccc ggggacttca agcaggagga ggtgggtggg 2940
133 gtgaaggagg aggcaggtgg gctgctgcag tgccccgagg tggccaaggc tgaccggtgg 3000
134 ctggaggaca gccggcactg ctgttccacc gccgacttcg gggacctccc actgctgcca 3060
135 cccaccagca ggaaggagga cctggaagct gaggaggagt actcctccct atgtgagctc 3120
136 ctgggcagcc ccgagcagag gcctggcatg caggacccgc tgtcacccaa ggccccactc 3180
137 atctgcacca aggaggaggt ggaggaggtg ctggactcca aggccggctg gggctctccg 3240
138 tgccacctct caggggagtc cgtcatcctg ctgggcccta cagtgggcac cgagtcaaag 3300
139 gtccagagct ggtttgagtc ctctctgtca cacatgaagc caggtgaaga ggggcctgat 3360
140 ggggagcgag ctccagggga ttccaccacc tcggacgcct ctctggccca gaagcccaac 3420
141 aagcctgctg tgcccgaggc gcccatcgca aagaaagagc ctgtgccacg gggcaaaagc 3480
142 ttacggagcc gtcgggtgca ccgggggctg cccgaggccg aggactcccc atgcagggca 3540
143 ccagtgctgc ccaaagacct cttgctccct gaatcctgca cagggccccc ccagggacag 3600
144 atggaagggg ctggagcccc aggccggggg gcctcggaag ggctccccag gatgtgtact 3660
145 cgttctctca cggccctgag tgagccccgc acgccccggac ccccaggcct gaccaccacc 3720
146 cctgcacccc cagacaaact ggggggcaag cagcgagccg ccttcaagtc gggcaagcgg 3780
147 gtggggaagc cctcacccaa ggctgcctcc agccccagca acccggccgc cctgcctgtg 3840
148 gcctccgaca gcagcccgat gggctccaag accaaggaga cagactcacc cagcacgcct 3900
149 ggcaaggacc agcgctccat gatccttcgg tcacgcacca aaacccagga gatcttccac 3960
150 tccaagcggc ggaggccctc tgagggccgg ctccccaact gccgtgccac caagaagctc 4020
151 ctcgacaaca gccacttgcc cgccacattc aaggtctcca gcagcccccа gaaggagggc 4080
152 agggtgagcc agcgggcaag ggtccccaaa cctggtgcag gcagcaagct ctctgaccgg 4140
153 ccctccatg cgctcaaaag gaagtcggcc ttcatggcgc cggtccccac caagaagcgg 4200
154 aacctggtct tgcggcacgg cagcagcagc agcagcaacg ccagtgcaat gggggagatg 4260
155 ggaaggagga gaggcctgag ggttccccca ccctcttcaa gaggatgtct tctcccaaga 4320
156 aagccaagcc caccaagggc aatggcgagc ctgccacaaa gctcccaccc ccggagaccc 4380
157 ccattcctgc ctcaagctcg cctctcggca gccttccagg gggccatgaa gaccaaggtg 4440
158 ctgccacccc ggaagggccg gggcctgaag ctggaagcca tcgtgcagaa gatcacctcg 4500
159 cccagcctca agaagttcgc atgtaaagcg ccaggggcct ctcctggtaa tcctctgagc 4560
160 ccatcccttt ccgacaaaga ccgtgggctc aagggtgctg ggggcagccc agtgggggtg 4620
161 gaagaaggcc tggtaaatgt gggcaccggg cagaagctcc caacttctgg ggctgatccg 4680
162 ttatgcagaa atccaaccaa cagatcctta aaaggcaaac tcatgaacag taagaaactg 4740
163 tcttctactg actgtttcaa aaccgaggcc ttcacatccc cggaggccct gcagcctggg 4800
164 gggactgccc tggcgcctaa gaagaggagc cggaaaggcc gggcaggggc ccatggactc 4860
165 tccaaaggcc cgctggagaa gcggccctat cttggcccgg ctctgctcct gactccccga 4920
166 gacagggcca gtggcacaca aggggccagt gaggacaact ctggtggagg aggcaagaag 4980
167 ccaaagatgg aggagctggg ccctgcctcc cagccccccgg agggcaggcc ctgccagccc 5040
168 cagacaaggg cacagaaaca gccaggccac accaactaca gcagctattc caagcggaag 5100
169 cgcctcactc ggggccgggc caagaacacc acctcttcac cctgtaaggg gcgtgccaag 5160
170 cgacgacgac aacagcaggt gctgcccctg gatcccgcag agcctgaaat ccgcctcaag 5220

RAW SEQUENCE LISTING DATE: 10/27/2000
PATENT APPLICATION: US/09/508,821 TIME: 13:38:12

Input Set : A:\01294811.app
Output Set: N:\CRF3\10272000\I508821.raw

171 tacatttcct cttgcaagcg gctgaggtca gacagccgga cccccgcctt ctcacccttc 5280
172 gtgcgggtgg agaagcgaga cgcgttcacc accatatgca ctgttgtcaa ctcccctgga 5340
173 gatgcgccca agccccacag gaagccttcc tcctctgcct cctcttcctc atcctcgtcc 5400
174 tcgttctcct tggatgcagc cggggcctcc ctggccacac tccctggagg ctccatcctg 5460
175 cagccgcggc cctccttgcc cctctcctcc acgatgcact tggggcctgt ggtttccaag 5520
176 gccctgagta cctcttgcct tgtttgctgc ctctgccaaa acccggccaa cttcaaggac 5580
177 cttggggacc tctgtgggcc ctactaccct gaacactgcc tccccaaaaa gaagccaaaa 5640
178 ctcaaggaga aggtgcggcc agaaggcacc tgtgaggagg cctcgctgcc gcttgagaga 5700
179 acactcaaag gtcccgagtg tgcagctgcc gccactgccg ggaagccccc caggtgacgg 5760
180 cccagctgac ccggccaagc agggcccact gcgcaccagt gcccggggcc tgtcccggag 5820
181 gctgcagagc tgctactgct gtgatggccg ggaggatggg ggcgaggagg cagccccagc 5880
182 cgacaagggt cgcaaacatg agtgcagcaa ggaggctccg gcagagcccg gcggggaggc 5940
183 ccaggagcac tgggtgcatg aggcctgtgc cgtgtggacc ggcggcgtct acctggtggc 6000
184 cgggaagctc tttgggctgc ag 6022
187 <210> SEQ ID NO: 6
188 <211> LENGTH: 22
189 <212> TYPE: DNA
190 <213> ORGANISM: Artificial Sequence
192 <220> FEATURE:
193 <223> OTHER INFORMATION: Description of Artificial Sequence:
194 oligonucleotide
196 <400> SEQUENCE: 6
197 tggccttgct gcccgtagtg ct 22

VERIFICATION SUMMARY DATE: 10/27/2000
PATENT APPLICATION: US/09/508,821 TIME: 13:38:13

Input Set : A:\01294811.app
Output Set: N:\CRF3\10272000\I508821.raw

L:44 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:2
L:44 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:2
L:44 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:2
L:44 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:2
L:44 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:2